

**WHAT IS CLAIMED IS:**

1. A method suitable for producing a consensus classification of organisms using the data derived from two or more experiments performed on said organisms or samples thereof comprising the steps of:
  - i) obtaining similarity matrices from the said data,
  - ii) producing a composite similarity matrix that is a function of said similarity matrices, and
  - iii) producing a consensus classification from said composite similarity matrix.
2. A method according to claim 1 wherein the function of step ii) comprises averaging the corresponding elements of said similarity matrices.
3. A method according to claim 2 wherein each similarity matrix is weighted according to the number of experimental characters used to calculate said matrix, to arrive at the average.
4. A method according to claim 2 wherein each similarity matrix is weighted by a user defined value to arrive at the average.
5. A method according to claim 2, wherein said experiments produce product size or retention time results, and wherein the each element of each similarity matrix is weighted according to the number of bands or features associated with that element, to arrive at the average.
6. A method according to claim 5 wherein said experiments are any of electrophoresis, high performance liquid chromatography, gas chromatography, capillary electrophoresis, chromatography, thin-layer chromatography, and/or mass spectrometry.
7. A method according to claim 1 wherein the function of step ii) comprises the steps of:
  - a) linearizing said similarity data matrices, and
  - b) averaging the corresponding elements of said linearized similarity matrices of step a)

8. A method according to claim 7 wherein step a) comprises the minimization of equations:

$$\sum_{i=1}^p \sum_{j=1}^{i-1} (\hat{d}_{k,ij} - f_k(D_{ij}))^2, \forall k$$

$$\sum_{i=1}^p \sum_{j=1}^{i-1} (D_{ij} - g_k(\hat{d}_{k,ij}))^2, \forall k$$

wherein  $p$  is the number of organisms, samples or genotypes, wherein each technique  $k$  results in a matrix of pair-wise distance values, so that the distance value obtained between organism

$i$  and  $j$  from technique  $k$  is given by  $d_{k,ij}$ , wherein  $\hat{d}_{k,ij} = \frac{d_{k,ij}}{S_k}$  with

$$S_k = \frac{2}{(p-1)(p-2)} \sum_{i=1}^p \sum_{j=1}^{i-1} d_{k,ij}, \text{ wherein the consensus distance matrix } D_{ij} \text{ is considered as the}$$

unknown true universal distance scale and wherein the goal is to search the consensus distances  $D_{ij}$ ,  $g_k$  and  $f_k$  so that  $\hat{d}_{k,ij} \cong f_k(D_{ij})$  and  $D_{ij} \cong g_k(\hat{d}_{k,ij})$  hold as true as possible.

9. An apparatus suitable for performing the methods according to claims 1 to 8.

10. A computer program comprising a computing routine, stored on a computer readable medium suitable for producing a consensus classification of organisms using the data derived from two or more experiments performed on said organisms or samples thereof according to the methods of claims 1 to 8.

11. A device suitable for producing a consensus classification of organisms using the data derived from two or more experiments performed on said organisms or samples thereof according to the methods of claims 1 to 8.